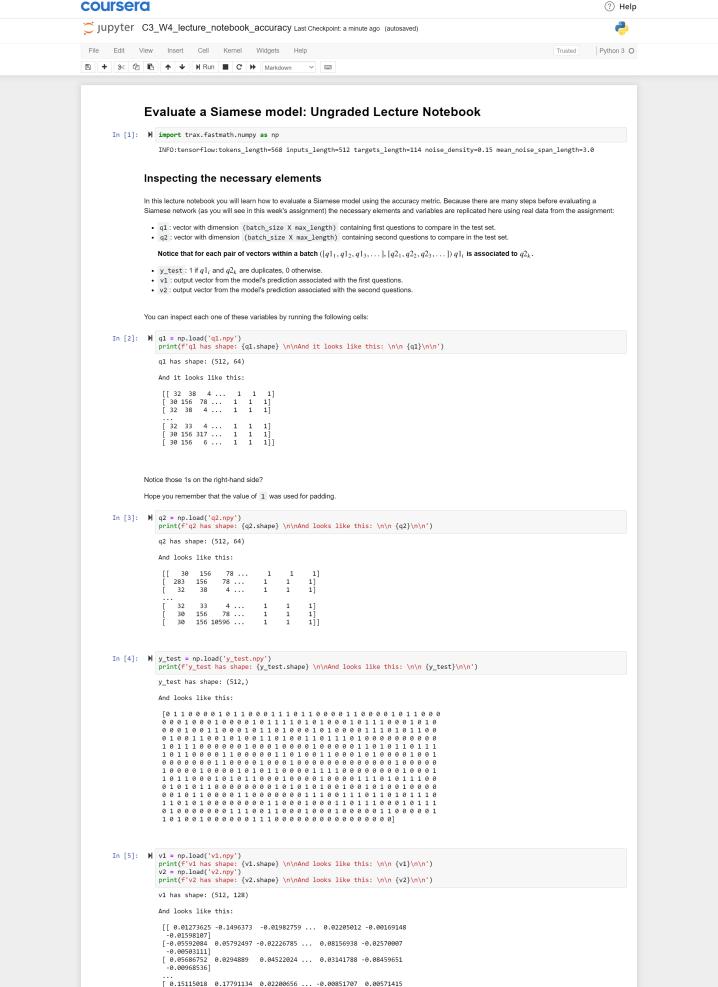
coursera



-0.00431194]

Calculating the accuracy

You will calculate the accuracy by iterating over the test set and checking if the model predicts right or wrong.

The first step is to set the accuracy to zero:

```
In [6]: ► accuracy = 0
```

You will also need the batch size and the threshold that determines if two questions are the same or not.

Note: A higher threshold means that only very similar questions will be considered as the same question.

```
In [7]: N batch_size = 512 # Note: The max it can be is y_test.shape[0] i.e all the samples in test data threshold = 0.7 # You can play around with threshold and then see the change in accuracy.
```

In the assignment you will iterate over multiple batches of data but since this is a simplified version only one batch is provided.

Note: Be careful with the indices when slicing the test data in the assignment!

The process is pretty straightforward:

- Iterate over each one of the elements in the batch
- Compute the cosine similarity between the predictions
 - For computing the cosine similarity, the two output vectors should have been normalized using L2 normalization meaning their magnitude will be 1. This has been taken care off by the Siamese network you will build in the assignment. Hence the cosine similarity here is just dot product between two vectors. You can check by implementing the usual cosine similarity formula and check if this holds or not.
- Determine if this value is greater than the threshold (If it is, consider the two questions as the same and return 1 else 0)
- Compare against the actual target and if the prediction matches, add 1 to the accuracy (increment the correct prediction counter)
- Divide the accuracy by the number of processed elements

```
In [8]: N for j in range(batch_size): # Iterate over each one of the elements in the batch

d = np.dot(v1[j],v2[j]) # Compute the cosine similarity between the predictions as L2 normalized, ||v1[j]||==||v2|
res = d > threshold # Determine if this value is greater than the threshold (if it is consider the two questic
accuracy += (y_test[j] == res) # Compare against the actual target and if the prediction matches, add 1 to the accuracy

accuracy = accuracy / batch_size # Divide the accuracy by the number of processed elements

In [9]: N print(f'The accuracy of the model is: {accuracy}')
```

The accuracy of the model is: 0.7421875

Congratulations on finishing this lecture notebook!

Now you should have a clearer understanding of how to evaluate your Siamese language models using the accuracy metric.

Keep it up!